

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
21. Juli 2005 (21.07.2005)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2005/066935 A2

(51) Internationale Patentklassifikation⁷: **G10L 17/00**

(21) Internationales Aktenzeichen: PCT/EP2004/014799

(22) Internationales Anmeldedatum:
29. Dezember 2004 (29.12.2004)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
103 61 850.3 31. Dezember 2003 (31.12.2003) DE

(71) Anmelder und

(72) Erfinder: **KRESS, Markus** [DE/DE]; Waldhornstrasse
52, 68199 Mannheim (DE).

(74) Anwälte: **BERNHARDT, Reinhold** usw.; Kobenhütten-
weg 43, 66123 Saarbrücken (DE).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,

MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,
GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,
ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU,
TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL,
PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Erklärungen gemäß Regel 4.17:

- hinsichtlich der Berechtigung des Anmelders, ein Patent zu
beantragen und zu erhalten (Regel 4.17 Ziffer ii) für alle
Bestimmungsstaaten
- Erfindererklärung (Regel 4.17 Ziffer iv) nur für US

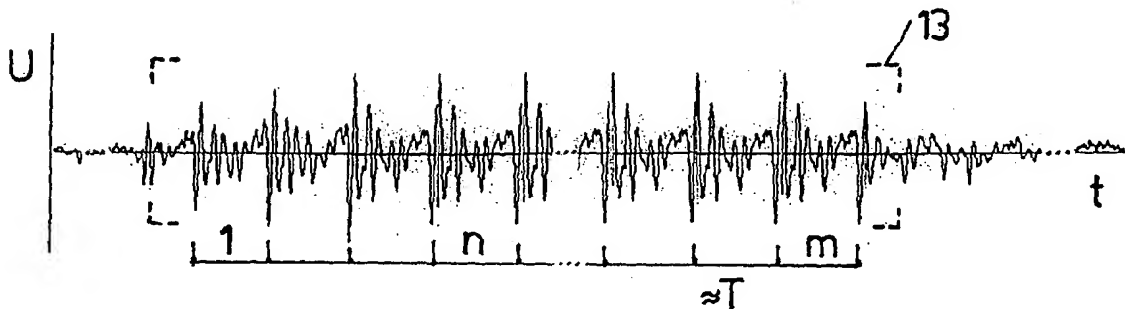
Veröffentlicht:

- ohne internationalen Recherchenbericht und erneut zu ver-
öffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Ab-
kürzungen wird auf die Erklärungen ("Guidance Notes on Co-
des and Abbreviations") am Anfang jeder regulären Ausgabe der
PCT-Gazette verwiesen.

(54) Title: METHOD FOR IDENTIFYING PEOPLE

(54) Bezeichnung: VERFAHREN ZUR IDENTIFIZIERUNG VON PERSONEN



(57) **Abstract:** The invention relates to a method for identifying people, whereby a person is identified by comparing an electric signal derived from a sound produced by the person with a stored signal of the above kind. The invention is characterized in that the signals to be compared are derived from the subphonemic range of sound production. The signal especially refers to a quasiperiod of a vowel or a semivowel.

(57) **Zusammenfassung:** Die Erfindung betrifft ein Verfahren zur Identifizierung von Personen, bei welchem eine Person durch Vergleich eines von einer bestimmten Lautäußerung der Person abgeleiteten elektrischen Signals mit einem gespeicherten solchen Signal identifiziert wird. Gemäß der Erfindung werden die zu vergleichenden Signale aus einem subphonemen Bereich der Lautäußerung abgeleitet. Insbesondere bezieht sich das Signal auf eine Quasiperioden eines Vokals oder eines Halbvokals.

WO 2005/066935 A2

Beschreibung:

Markus Kreß, D-68199 Mannheim (Deutschland)

5

Verfahren zur Identifizierung von Personen

Die Erfindung betrifft ein Verfahren zur Identifizierung von Personen, bei dem eine Person durch Vergleich eines von einer bestimmten Lautäußerung der Person abgeleiteten elektrischen Signals mit einem gespeicherten solchen Signal identifiziert wird.

Aus der EP 0 896 711 B1 und der DE 100 42 571 C2 bekannte solche Verfahren zur Identifizierung von Personen anhand ihrer Stimme verwenden zum Vergleich ein Signal, dass der Gesamtheit der Lautäußerung oder einer daraus ausgewählten Lautfolge entspricht. In diesen Signalen enthaltene individuelle Merkmale lassen eine mehr oder weniger sichere Personenidentifizierung zu.

Je nach Anzahl gespeicherter Vergleichssignale, d.h. je nach der Größe des Personenkreises, aus dem einzelne Personen identifiziert werden sollen, kann ein Identifizierungsvorgang nach einem solchen Verfahren verhältnismäßig langwierig und das Verfahren z.B. für die Kontrolle der Zugangsbefugnis zu einem größeren Betrieb oder einer größeren Institution ungeeignet sein.

Der Erfindung liegt die Aufgabe zugrunde, ein neues Verfahren der eingangs erwähnten Art zu schaffen, dass bei höherer Identifizierungssicherheit eine schnellere Personenidentifizierung als die bekannten derartigen Verfahren ermöglicht.

Dass diese Aufgabe lösende Verfahren nach der Erfindung ist dadurch gekennzeichnet, dass die zu vergleichenden Signale aus einem subphonemen Bereich der Lautäußerung abgeleitet werden.

BESTÄTIGUNGSKOPIE

Der Erfindung liegt die Erkenntnis zugrunde, dass bereits ein Signal, das von einem kurzen, als solchem nicht hörbaren Ausschnitt aus der Lautäußerung bzw. aus einem dieser entsprechenden elektrischen Gesamtsignal abgeleitet ist, genügend viele, für ein Individuum charakteristische Merkmale aufweist, um eine Identifizierung durchführen zu können. Vorteilhaft ist aufgrund der Kürze des Signals der Umfang der bei einem Identifizierungsvorgang zu verarbeitenden Daten gegenüber den bekannten Verfahren erheblich geringer, wodurch sich der Identifizierungsvorgang merklich verkürzt. Zudem treten die individuellen Merkmale in dem kürzeren Vergleichssignal deutlicher hervor, während diese in einem elektrischen Signal, das einer längeren Lautfolge entspricht, stärker „verwischt“ sind. Folglich wird durch die Erfindung auch die Identifizierungssicherheit erhöht. Fehlerhafte Feststellungen einer Übereinstimmung oder Nichtübereinstimmung von Vergleichssignalen sind nahezu ausgeschlossen.

Vorzugsweise wird in einem ersten Schritt zur Ableitung der zu vergleichenden Signale ein der gesamten Lautäußerung entsprechendes elektrisches Ausgangssignal eines elektroakustischen Wandlers einer Lautstärkenormierung unterzogen. Vorteilhaft werden dadurch Signalunterschiede, die auf keinerlei individueller Prägung beruhen, ausgefiltert. Die Lautstärkenormierung kann bereits in einer Mikrofonerätteeinheit erfolgen, die an einen Computer mit einem Mikrofoneingang anschließbar ist.

Im Computer wird das Ausgangssignal digitalisiert und zweckmäßig eine das Ausgangssignal approximierende Fourierreihe gebildet, die der weiteren Signalverarbeitung im Computer zugrundegelegt werden kann.

In einer bevorzugten Ausführungsform der Erfindung wird in dem digitalisierten Ausgangssignal des elektroakustischen Wandlers, das der Lautäußerung entspricht, ein quasiperiodischer Bereich des Signals ermittelt. Ein quasiperiodischer Bereich ist stets dann vorhanden, wenn die Lautäußerung einen Vokal oder Halbvokal enthält.

Während aus dem quasiperiodischen, z.B. dem Buchstaben a entsprechenden, Bereich irgendein subphonemer Teilbereich auswählbar ist, wird in der bevorzugten Ausführungsform der Erfindung zur Bildung eines Vergleichssignals oder zur Bildung mehrerer Vergleichssignale jeweils nur eine einzige Quasiperiode ausgewählt.

Zweckmäßig handelt es sich dabei um eine bestimmte Quasiperiode n aus dem quasiperiodischen Bereich 1 bis m . Nicht individuelle Signalprägungen, die lediglich auf unterschiedlichen Stellungen der Vergleichsperiode innerhalb des quasiperiodischen Bereiches beruhen, werden so ausgefiltert.

5

In weiterer vorteilhafter Ausgestaltung der Erfindung wird die ausgewählte Quasiperiode einer Längennormierung unterzogen, d.h. auf eine Normlänge T gedehnt oder gestaucht. Schwankungen der Periodenlänge innerhalb der Quasiperiodizität und insbesondere von der Stimmhöhe abhängige Periodendauerunterschiede werden so ausgeglichen und individuelle Prägungen des Signals genauer bestimmten Zeitpunkten innerhalb der Periode T zugeordnet. Sie treten im Vergleich dadurch genauer hervor.

In weiterer Ausgestaltung der Erfindung wird aus der ausgewählten Quasiperiode und einer über eine Vielzahl von Personen gemittelten solchen Quasiperiode ein Quotientensignal als Vergleichssignal gebildet.

Ein solches Quotientensignal ist also auf ein Signal bezogen, dass nur wenig individuell geprägt ist. Dementsprechend treten in dem Quotientensignal individuelle Prägungen umso stärker hervor.

Ferner werden in der bevorzugten Ausführungsform der Erfindung mehrere, z. B. drei, zu speichernder Vergleichssignale gebildet, indem die Lautäußerung in unterschiedlichen Stimmhöhen aufgezeichnet und verarbeitet wird. Bei der Identifizierung erfolgt eine Interpolation bzw. es wird durch Interpolation eine Kurvenschar von gespeicherten Vergleichssignalen gebildet.

Das Identifikationsverfahren kann Bestandteil eines Spracherkennungsprogramms, die Vergleichssignale können Bausteine eines Syntheseprogramms für Sprache sein.

Die Erfindung soll nun anhand eines Ausführungsbeispiels und der beiliegenden, sich auf dieses Ausführungsbeispiel beziehenden Zeichnungen näher erläutert werden. Es zeigen:

35

Fig. 1 eine Schemadarstellung einer nach dem erfindungsgemäßen Verfahren arbeitenden Identifizierungsvorrichtung, und

Fig. 2 ein einer Lautäußerung entsprechendes elektrisches Signal, aus dem ein zur Personenidentifizierung geeignetes Vergleichssignal gemäß der Erfindung ableitbar ist.

5 In Fig. 1 weist das Bezugszeichen 1 auf einen elektroakustischen Wandler hin, dem eine Einrichtung 2 nachgeschaltet ist, welche eine Lautstärkenormierung durchführt. Der elektroakustische Wandler 1 und die Normierungseinrichtung 2 sind zu einer Geräteeinheit 3 zusammengefasst, welche an einen Mikrofoneingang eines Computers 4 angeschlossen ist.

10

Der Computer 4 enthält durch Hard- und Software gebildete Einrichtungen 5 bis 12.

15 Eine Digitalisierungseinrichtung 5 empfängt das Ausgangssignal der Geräteeinheit 3. Das durch die Einrichtung 5 digitalisierte Signal gelangt in eine Einrichtung 6, in welcher eine das Signal approximierende Fourierreihe gebildet und der weiteren Signalverarbeitung zugrundegelegt wird.

20 Eine Einrichtung 7 ermittelt einen quasiperiodischen Bereich des Signals, aus welchem eine nachgeordnete Einrichtung 8 mindestens eine bestimmte Quasiperiode auswählt. Es können Quasiperioden auch aus mehreren ermittelten quasiperiodischen Bereichen ausgewählt werden.

25 In einer nachgeordneten Einrichtung 9 erfolgt eine Bearbeitung der ausgewählten Quasiperiode, z.B. eine zeitliche Streckung oder Stauchung auf eine Normdauer.

Je nachdem, ob ein Vergleichssignal gespeichert oder eine Person identifiziert werden soll, wird die bearbeitete Quasiperiode als Vergleichssignal einer Speichereinrichtung 10 oder einer Vergleichereinrichtung 12 zugeführt.

30

In der Vergleichereinrichtung 12 wird die bearbeitete Quasiperiode mit solchen gespeicherten Signalen einer Vielzahl von Personen verglichen und durch Feststellung der Übereinstimmung mit einem der gespeicherten Signale eine Person identifiziert.

35

Eine Mittelungseinrichtung 14 bildet aus den für die Vielzahl von Personen gespeicherten Signalen ein gemittelttes Signal, dass in der Speichereinrichtung 10 gespeichert und der Bearbeitungseinrichtung 9 zugeführt werden kann.

Im folgenden wird ein Identifizierungsvorgang anhand von Fig. 2 noch näher erläutert.

- 5 Eine zu identifizierende Person, von welcher in der Speichereinrichtung 10 ein Vergleichssignal gespeichert ist, spricht ein vorgegebenes Wort, beispielsweise das Wort „Mama“. Die Geräteeinheit 3 bildet aus einem entsprechenden akustischen Signal 14 ein Lautstärkenormiertes Signal $U(t)$. Ein den ersten Vokal „a“ des Wortes „Mama“ betreffender Teil dieses Signals ist in Fig. 2 dargestellt.
- 10 Das gesamte, dem Wort „Mama“ entsprechende Lautstärkenormiertes Signal $U(t)$ wird durch die Einrichtung 5 digitalisiert und die Funktion $U(t)$ anschließend in der Einrichtung 6 durch eine Fourierreihe dargestellt. Weitere Signalbearbeitungen erfolgen auf der Grundlage dieser Fourierreihe.
- 15 Im nächsten Verarbeitungsschritt ermittelt die Einrichtung 7 mit Hilfe eines zeitlich variablen Beobachtungsfensters 13 in dem Gesamtsignal $U(t)$ einen ersten quasiperiodischen Bereich mit Quasiperioden 1 bis m und wählt aus dem Bereich mindestens eine Quasiperiode n aus.
- 20 Da die Dauer der Quasiperioden etwas schwankt und darüber hinaus von der jeweiligen Stimmenhöhe abhängt, erfolgt in der Bearbeitungseinrichtung 9 eine Streckung oder Stauchung der ausgewählten Periode n auf eine Normdauer T . Ferner wird in der Einrichtung 9 ein Quotientensignal aus der gestreckten oder gestauchten Periode n und einem durch die Einrichtung 11 erzeugten und in der
- 25 Speichereinrichtung 10 gespeicherten Signal gebildet. Letzteres stellt eine Mittelung der Signale einer Vielzahl von Personen dar. In diesem Quotientensignal treten individuelle Besonderheiten deutlich hervor. Über dieses Quotientensignal hinaus kann ferner eine Quotientenbildung mit einem unter besonderen emotionalen Bedingungen aufgenommen Vergleichssignal erfolgen.
- 30 Das durch die Bearbeitungseinrichtung 9 bearbeitete Vergleichssignal wird, falls es sich um eine Probeaufnahme von einer in den Kreis der zu identifizierenden Personen aufzunehmenden Person handelt, in der Einrichtung 10 abgespeichert, wobei im Falle einer solchen Probeaufnahme mehrere, z. B. drei, Vergleichssignale
- 35 gebildet werden, nämlich für drei unterschiedliche Stimmenhöhen, in denen das Wort „Mama“ auszusprechen ist. Im Identifizierungsfall wird das betreffende Signal der Vergleichereinrichtung 12 zugeführt, in welcher ein Vergleich mit allen in der Einrichtung 10 gespeicherten Vergleichssignalen erfolgt. Wird eine Überein-

stimmung mit einem der gespeicherten Signale festgestellt, so ist die Person als dem Personenkreis zugehörig identifiziert.

5

10

15

20

25

30

35

Patentansprüche:

1. Verfahren zur Identifizierung von Personen, bei dem eine Person durch Vergleich eines von einer bestimmten Lautäußerung der Person abgeleiteten elektrischen Signals mit einem gespeicherten solchen Signal identifiziert wird,
5 dadurch gekennzeichnet,
dass die zu vergleichenden Signale aus einem subphonemen Bereich der Lautäußerung abgeleitet werden.
- 10 2. Verfahren nach Anspruch 1,
dadurch gekennzeichnet,
dass in einem ersten Schritt zur Ableitung der Signale ein der gesamten Lautäußerung entsprechendes elektrisches Ausgangssignal eines elektroakustischen Wandlers (1) einer Lautstärkenormierung unterzogen wird.
- 15 3. Verfahren nach Anspruch 1 oder 2,
dadurch gekennzeichnet,
dass eine ein der gesamten Lautäußerung entsprechendes Ausgangssignal approximierende Fourierreihe gebildet wird.
- 20 4. Verfahren nach Anspruch 2 oder 3,
dadurch gekennzeichnet,
dass zur Ableitung der zu vergleichenden Signale mindestens ein quasi-periodischer Bereich des Ausgangssignals ermittelt wird.
- 25 5. Verfahren nach Anspruch 4,
dadurch gekennzeichnet,
dass zur Ableitung der zu vergleichenden Signale aus dem ermittelten quasi-periodischen Bereich eine einzige Quasiperiode oder mehrere Quasiperioden
30 ausgewählt wird/werden.
6. Verfahren nach Anspruch 5,
dadurch gekennzeichnet,
dass eine in Bezug auf ihre Stellung in dem quasiperiodischen Bereich (1 bis m)
35 bestimmte Quasiperiode (n) ausgewählt wird.

7. Verfahren nach Anspruch 5 oder 6,
dadurch gekennzeichnet,
dass die ausgewählte Quasiperiode einer Längennormierung unterzogen wird.
- 5 8. Verfahren nach einem der Ansprüche 5 bis 7,
dadurch gekennzeichnet,
dass aus der ausgewählten Quasiperiode und einer für eine durchschnittliche
Stimme maßgebenden Quasiperiode ein Quotientensignal gebildet wird.
- 10 9. Verfahren nach einem der Ansprüche 1 bis 5,
dadurch gekennzeichnet,
dass zu Bildung zu speichernder Vergleichssignale die Lautäußerung mehrfach
in unterschiedlichen Stimmhöhen aufgenommen und bei der Identifizierung
zwischen mehreren Vergleichssignalen interpoliert oder durch Interpolation
15 eine Kurvenschar von Vergleichssignalen gebildet wird.
10. Verfahren nach einem der Ansprüche 1 bis 9,
dadurch gekennzeichnet,
dass das Verfahren in ein Spracherkennungsprogramm integriert wird.
- 20 11. Verfahren nach einem der Ansprüche 1 bis 10,
dadurch gekennzeichnet,
dass die zu vergleichenden Signale als Bausteine eines Sprachsynthese-
programms dienen.

25

30

35

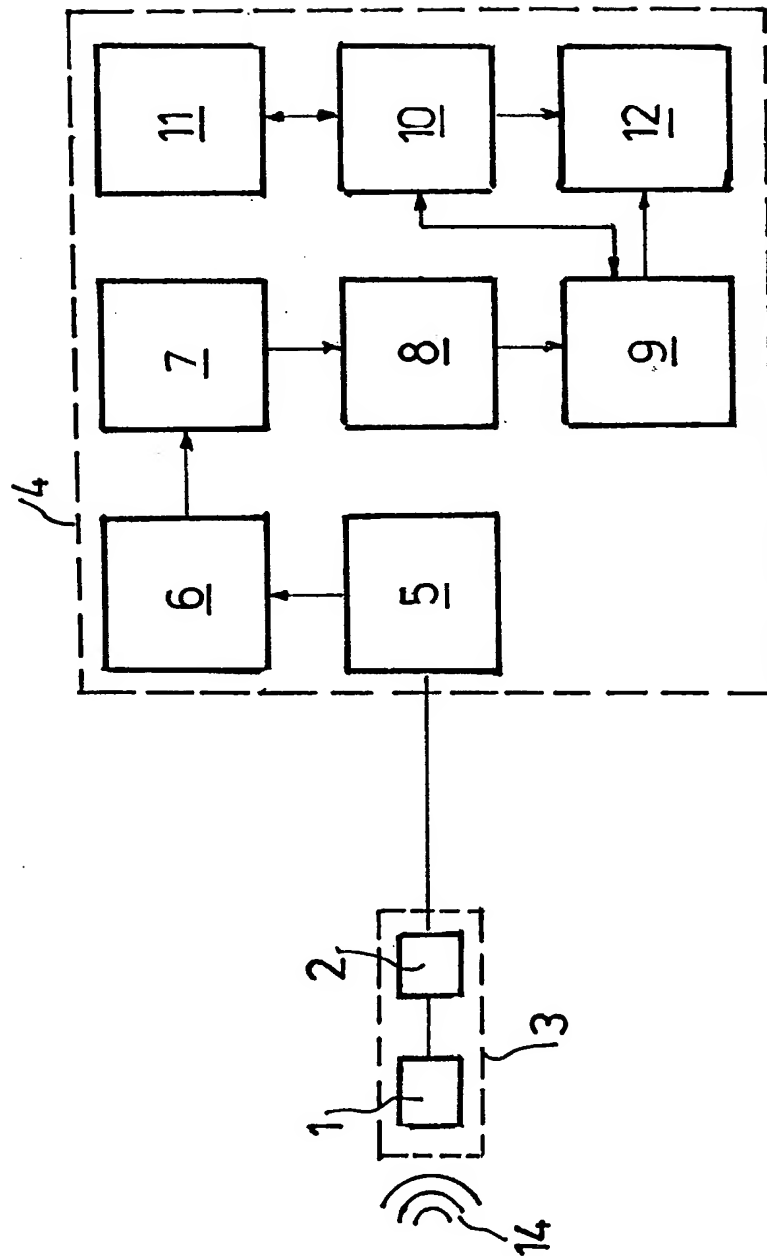


FIG.1

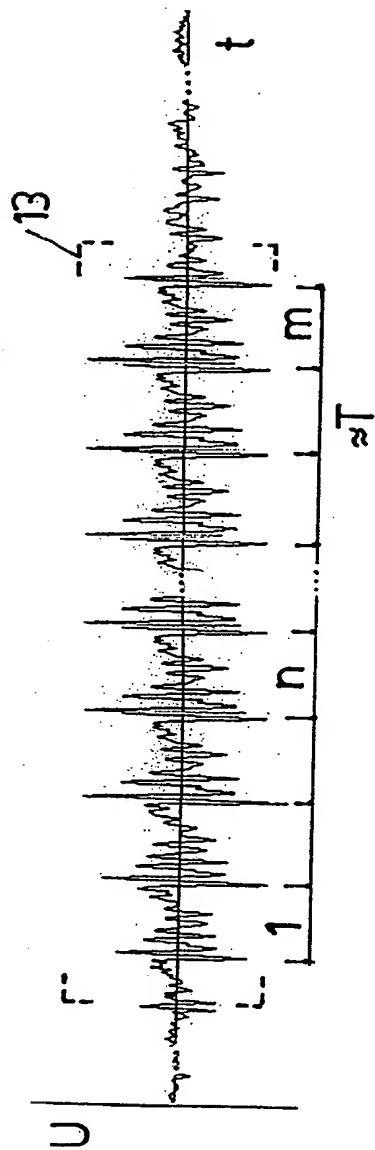


FIG.2